A RELIABILITY TEST METHOD FOR "ONE-SHOT" ITEMS

H. J. Langlie
Aeronutronic Division, Ford Motor Company

1. INTRODUCTION. As a result of many reliability problems which have plagued procuring agencies in the missile and space programs, increased emphasis has been placed on the development of improved reliability demonstration test methods. In this connection, the Army has advocated increased use of the "test-to-failure" concept to establish the existence of satisfactory margins of reliability with respect to critical factors. This philosophy has the advantage in that statistical statements can be made regarding reliability on the basis of relatively small samples. This paper will discuss the application of this general concept to a particular class of hardware.

In essence, the "test-to-failure" concept involves submitting a test specimen to an increasing environment or load stress until failure is detected. By observing the statistical behavior of the stresses at which failure occurs, the lower limit of stress below which the probability of failure is very small can be selected by using the mean and standard deviation of these data. Robert Lusser (reference 1) advocated the safety margin concept in interpreting these data. As is shown in Figure 1, the larger the value of k (which is the distance from the mean strength to the upper limit of the operating applied stress divided by the standard deviation of the data), the greater the reliability of the specimen with respect to the stress involved.

In establishing the reliability objectives for the Shillelagh Program, the Army Missile Command required that safety margins be demonstrated in the test laboratory for Shillelagh components with respect to critical environmental stress factors. Many of these components are of a "go-no-go" type such as thermal batteries, electrical relays, and other short-lived equipment items. In most cases, little information was available prior to test regarding the nature of the standard deviation of the distribution of strenghts for these parts. Furthermore, it was desired to perform a laboratory test involving a minimum number of samples. A review of attribute sensitivity testing techniques such as the Up and Down method (reference 2) and the Probit method (reference 3) indicated that these methods cannot be applied satisfactorily under the sample size and technical limitations imposed. As a result, a study was made to develop a method for selecting stress levels for testing which required no a'priori assumption regarding the standard deviation of the unknown strength distribution and could be performed satisfactorily with sample sizes of the order of fifteen or twenty.

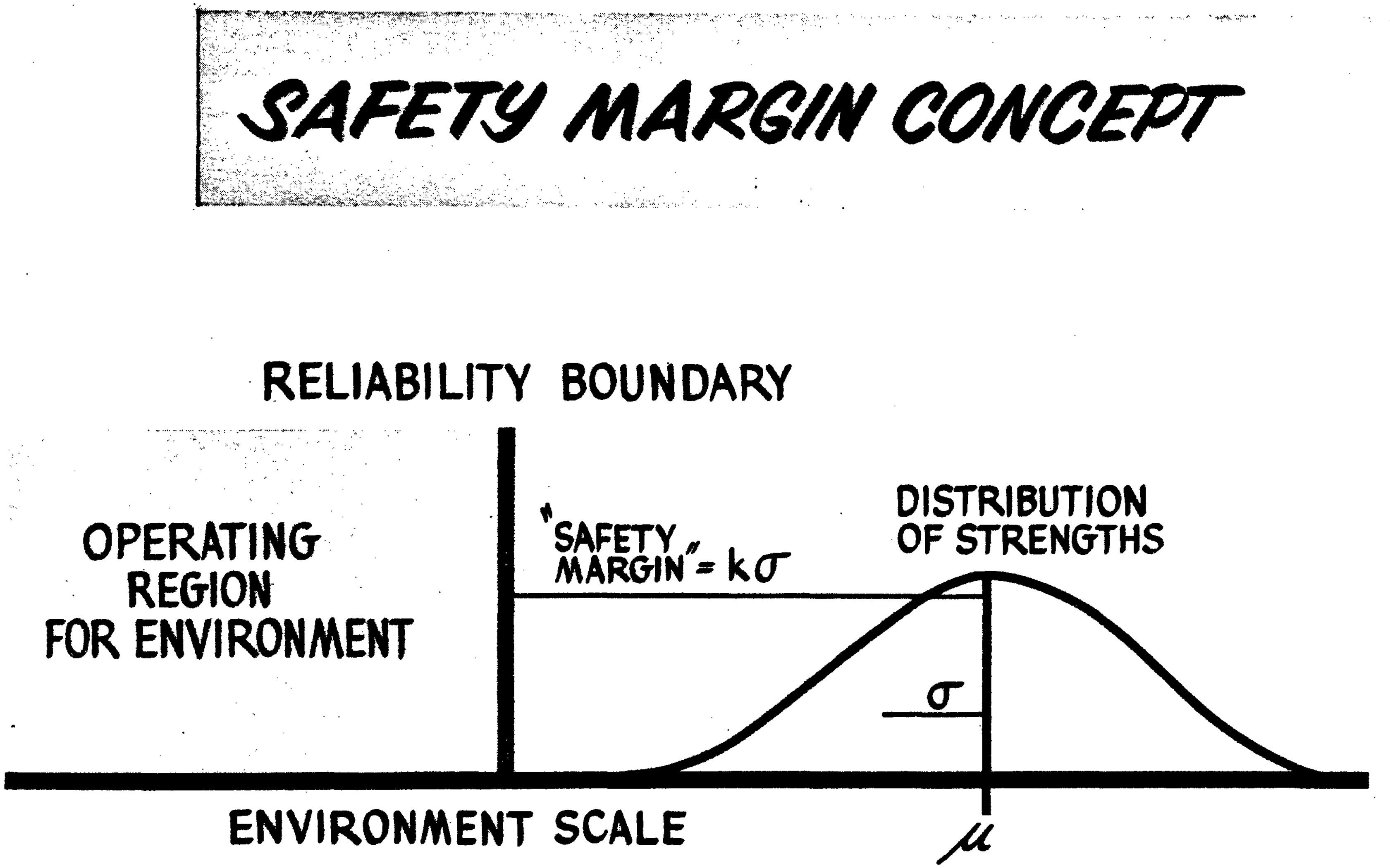

SAFETY MARGIN CONCEPT
RELIABILITY BOUNDARY
OPERATING REGION FOR ENVIRONMENT
"SAFETY MARGIN" = kσ
DISTRIBUTION OF STRENGTHS
σ
ENVIRONMENT SCALE
μ

FIGURE 1.

Detailed empirical investigations were performed, using Monte Carlo methods with a high speed digital computer, to develop a satisfactory algorithm for determining successive stress test levels as the experiment proceeds. Exact maximum likelihood equations were used to calculate the statistical estimates, μ_e, and σ_e, of the population parameters, μ and σ, of the distribution of part strengths. By repeated simulations of experiments for samples sizes ranging from ten to 900, empirical curves were obtained showing the variance in the calculated estimates vs sample size. After this study was complete, a technical report was prepared (reference 4). The paper today will discuss some significant results contained in this report.

Since the completion of the study, several applications have been made in reliability testing electrical and mechanical components. One such application is presented in this paper by way of illustration.

2. DISCUSSION. In order to proceed with the discussion of the test method, a specific definition is given regarding the terms "stress" and "strength" as follows:

> Stress is a test factor, such as environmental level or force level which is applied to the test specimen. Operational stresses represent the mix of environmental or load conditions that can be expected to be imposed on a typical specimen during its life. During the conduct of a "one-shot" test, the stress represents the applied test factor which is varied in magnitude from specimen to specimen in a systematic manner.
>
> Strength is a property ascribed to a specimen such that if the stress imposed on the part is greater than the strength of the part, the part will fail. Conversely, if the stress is less than the strength of the part, the part will not fail.

Failure in the sense used above is a general term referring to unsatisfactory completion of function, out of tolerance performance, breakage, or other evidence of malfunction. For each test attempt wherein a stress is applied to a specimen, there is associated an outcome which is a binary variable: success or failure.

It is assumed that, given a homogenous sample of replicate specimens, the part strengths are distributed normally with an unknown mean and

standard deviation. The purpose of the test method is to select stress levels in such a way as to generate outcomes which can then be used to calculate statistical estimates of the parameters of the distribution of part strengths.

3. PERFORMING THE TEST. In undertaking to perform the "one-shot" test, there are three steps to be taken which should be followed regardless of the nature of the application. They are:

1. Establish the criteria of failure, or acceptance.
2. Determine the test interval.
3. Select the stress levels.

The last of these three steps, selecting the stress levels, proceeds concurrent with the actual performance of the test. These three steps are discussed below.

3.1 Establishing Failure Criteria.

It is very important that painstaking care be given to the set of ground rules which will be followed for differentiating between a success or a failure. For purposes of reliability testing, this means that a careful enumeration should be made of all undesirable responses of a test specimen, such as particular modes of failure, out of tolerance performance, and any other mode of unacceptable product performance. These criteria are the very basis for establishing product assurance in the laboratory and should be reviewed and approved by all qualified parties having a technical interest in the product.

3.2 Determining the Test Interval.

In order to proceed with a generation of stress levels for testing purposes, it is necessary to choose a test interval which is used as a basis for the stress sequencing method. This interval should be selected large enough to include all possible ranges of strengths of the parts to be tested. This interval can be made conservatively large, since the "one-shot" method has been designed to cause the stress levels to be generated in the vicinity of interest (i.e., in the vicinity of the distribution of strengths) as the test proceeds. As a sample illustration, the range for a drop height test for glass containers designed to withstand say, a six inch drop, could be chosen to have a lower limit of zero and an upper limit of three feet. The method of analysis of the data is such that the particular choice of the

endpoints of the test interval do not have an appreciable effect on the results for sample sizes of fifteen or more. In the event that the test interval turns out, as the test proceeds, to be inappropriately chosen, then the stress levels will tend to converge towards one limit or the other. In such an event, particularly in the case of reliability testing, convergence towards the lower level is usually indicative of a totally unsatisfactory product, whereas convergence towards the upper limit can be shown to be statistically acceptable by use of the likelihood ratio test.

In Figure 2 there is represented the results of an actual "one-shot" test on thermal batteries to determine the reliability with regard to high temperature. In this instance, the batteries were designed to perform reliably at 145°F. On the basis of conservative engineering judgement and some limited development test data, the lower limit was selected to be 100°F (the level at which all thermal batteries would be expected to perform satisfactorily) and the higher limit was selected to be 350°F (the level at which <u>all</u> thermal batteries would be expected to fail).

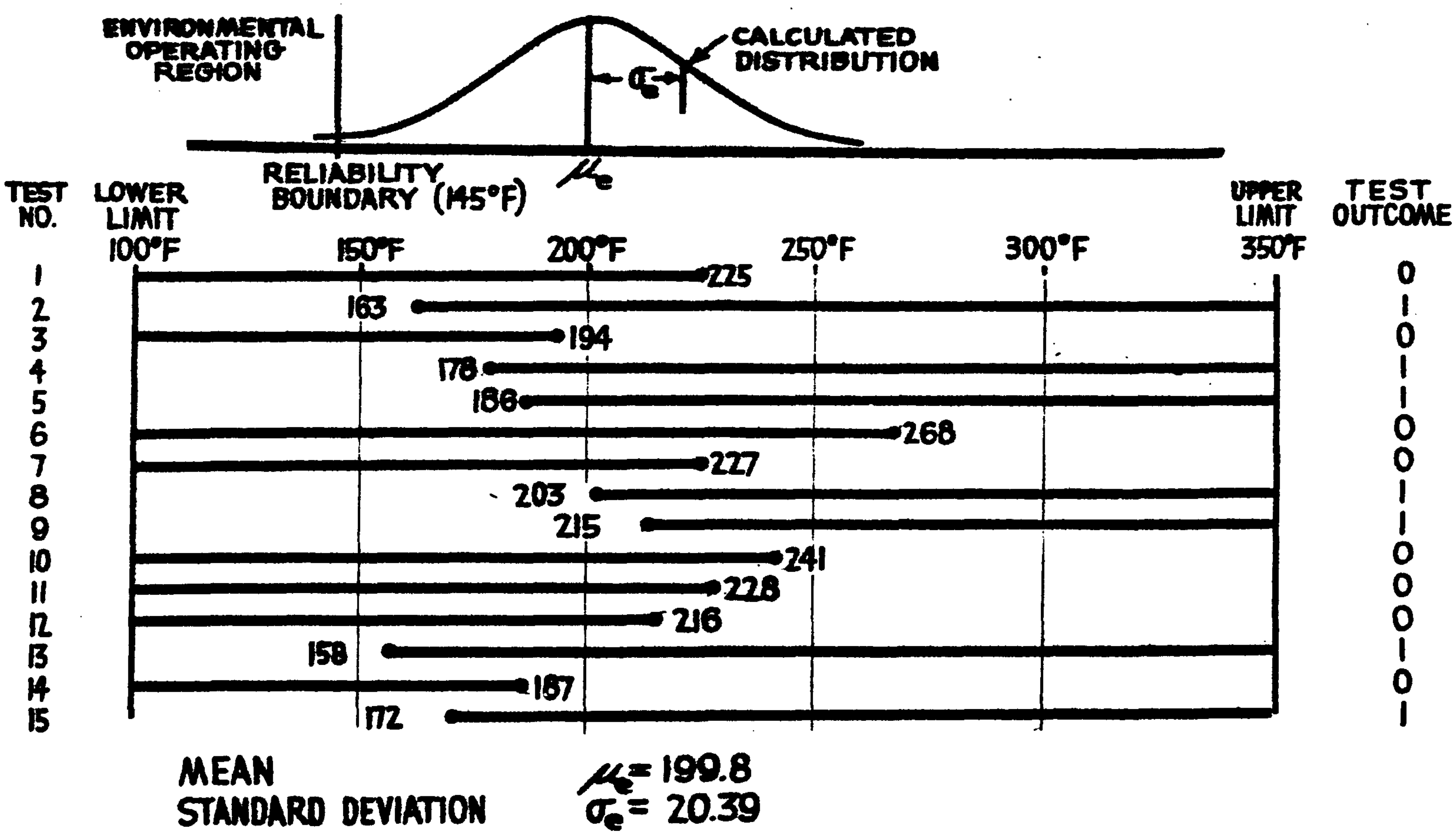
3.3 Selecting the Stress Levels.

Once the test interval and failure criteria have been established, the test commences by selecting the first stress level at the midpoint of the interval. After exposing the first specimen to this environmental level and activating it, a one or zero is recorded to indicate the outcome as a success or failure respectively (see Figure 2).

The general rule for obtaining the $(n+1)^{st}$ stress level, having completed n trials, is to work backward in the test sequence, starting at the n^{th} trial, until a previous trial (call is the p^{th} trial) is found such that there are as many successes as failures in the p^{th} through the n^{th} trials. The $(n+1)^{st}$ stress level is then obtained by averaging the n^{th} stress level with the p^{th} stress level. If there exists no previous stress level satisfying the requirement stated above, then the $(n+1)^{st}$ stress level is obtained by averaging the n^{th} stress level with the lower or upper stress limits of the test interval according to whether the n^{th} result was a failure or a success.

To illustrate, suppose it is desired to find the second stress level in Figure 2. Since there was only one previous observation (i.e., first unit failed) it is not possible to find a stress level where all intervening results even out. That is, the second stress level is obtained by averaging the first with the lower limit. To find the eighth stress level, it is observed that results from test 4 through 7 (i.e., the last four results) cancel each other out. Thus, the eighth stress level is obtained by averaging the fourth.

SAMPLE "ONE-SHOT" TEST



TEST-TO-FAILURE OF THERMAL BATTERIES IN TEMPERATURE

FIGURE 2.

stress level with the seventh.

As a final example, it is observed that after the twelfth test has been completed, there again exists no previous stress level for which the number of failures equals the number of successes. Since the twelfth test was a failure, the thirteenth stress level is obtained by averaging the twelfth stress level with the lower limit.

As an aid in identifying the important parameters of the test, the stress level is designated by the letter s and the outcome is designated by the letter u. The lower limit of the test interval is designated A and the upper limit is designated B. Upon the conclusion of the test, the stress values, $(s_1, s_2, \ldots, s_N)$, and the corresponding outcomes, $(u_1, u_2, \ldots, u_N)$, where N equals the test sample size, are used to perform a complete analysis for hardware reliability.

4. DERIVATION OF MAXIMUM LIKELIHOOD EQUATIONS. Consider the random sample $X = (x_1, x_2, \ldots, x_N)$ of N observations where the x_i are independent random variable from a Gaussian distribution, $g(x;\mu, \sigma)$, with mean μ and standard deviation σ. Consider also an N-dimensional vector $S = (s_1, s_2, \ldots s_N)$ where $A \leq s_i \leq B$. From this, construct a third vector $U = (u_1, u_2, \ldots, u_N)$ where

$$u_i = 1 \text{ if } s_i < x_i$$
$$u_i = 0 \text{ if } s_i \geq x_i$$

The variable x_i is called the strength of the i^{th} part; s_i is called the applied stress level for the i^{th} part, and u_i is called the outcome of the "test" on the i^{th} part. The outcome $u = 1$ is called a success (i. e., the applied stress was less than the part strength) and, conversely, $u = 0$ is called a failure.

The object of this development is to obtain formulas for calculating the estimates μ_e and σ_e of μ and σ, given only S and U. This will be accomplished by obtaining values of μ_e and σ_e which maximize the likelihood (i. e., probability) of obtaining the outcome U given S.

The probability of outcome u_i given s_i can be written

(1) $$p_i = \text{Prob}\left[u_i \middle| s_i\right] = u_i \int_{s_i}^{\infty} g(v;\mu,\sigma)dv + (1 - u_i)\int_{-\infty}^{s_i} g(v;\mu,\sigma)dv$$

The probability of outcome U can be written as the product of the probabilities of the individual u_i since the x_i are independent.

(2) $$P\left[U\right] = \prod_{i=1}^{N} p_i = L(\mu, \sigma)$$

The expression $P\left[U\right]$, when regarded as a function of the population parameters μ and σ, becomes the likelihood function for outcome U (ref. 5).

To find the values of μ and σ (now regarded as variables) which maximize L, we differentiate (2) with respect to μ and σ and solve the system

(3) $$\partial \ln L/\partial\mu = O$$ $$\partial \ln L/\partial\mu = O$$

where the logarithm of L is used to simplify the algebra.

Letting $t_i = (s_i - \mu)/\sigma$, $g_o(v) = (2\pi)^{-1/2} \exp(-v^2/2)$, the normalized Gaussian, and remembering that u_i can take on only values of O or 1, equation (1) can be re-written

(4) $$\ln p_i = u_i \ln\left[\int_{t_i}^{\infty} g_o(v)dv\right] + (1 - u_i)\ln\left[\int_{-\infty}^{t_i} g_o(v)dv\right]$$

The following definitions ($\overset{d}{=}$) and derivations will be helpful:

(5) $$G(t) \overset{d}{=} \int_{-\infty}^{t} g_o(v)\, dv$$

(6) $$dg_o(t)/dt = -tg_o(t)$$

(7) $$\partial t/\partial\mu = -1/\sigma \qquad (\sigma > O)$$

(8) $$\partial t/\partial\sigma = -t/\sigma \qquad (\sigma > O)$$

Since $\ln L = \Sigma \ln p_i$, equations (3) become

(9) $$\partial \ln L/\partial\mu = \Sigma\, \partial \ln p_i/\partial\mu = O$$

$$\partial \ell n L/\partial\sigma = \Sigma\, \partial \ln p_i/\partial\sigma = O$$

From (4)

$$(10) \qquad \partial \ln p_i/\partial\mu = \frac{-u_i g_o(t_i)}{1 - G(t_i)} \frac{\partial t_i}{\partial\mu} + \frac{(1-u_i)g_o(t_i)}{G(t_i)} \frac{\partial t_i}{\partial\mu}$$

$$= \frac{g_i}{\sigma}\left[\frac{u_i}{1-G_i} - \frac{1-u_i}{G_i}\right]$$

where the arguments (and subscript "o") have been omitted to simplify writing. Similarly,

$$(11) \qquad \partial \ln p_i/\partial\sigma = \frac{t_i g_i}{\sigma}\left[\frac{u_i}{1-G_i} - \frac{1-u_i}{G_i}\right]$$

Denoting by h_i the expression in brackets and eliminating the constant σ, equations (9) become

$$\left.\begin{aligned} \partial \ln L/\partial\mu &\overset{d}{=} p(\mu,\sigma) = \sum_{i=1}^{N} g_i h_i = 0 \\ \partial \ln L/\partial\sigma &\overset{d}{=} q(\mu,\sigma) = \sum_{i=1}^{N} t_i g_i h_i = 0 \end{aligned}\right\}$$

Equations (12) are valid only if the value of σ which satisfies the maximum likelihood equations is non-zero. A quick examination of the data can be made to determine if a non-zero σ is a maximum likelihood solution. If the maximum stress level at which a success occurred is greater than the minimum level at which a failure occurred, than a non-zero σ satisfies equations (12). If this statement is not true, then $\sigma = 0$ represents a maximum likelihood estimate for the standard deviation and the maximum likelihood estimate of the mean is a connected interval contained between the maximum failure stress level and the minimum success stress level which represent the lower and upper bounds of the interval respectively. The latter situation illustrates an outcome which in fact must be achieved if all of the part strengths were concentrated at a mass point within the above mentioned interval. The maximum likelihood corresponding to this outcome is one. Although unique estimates cannot be obtained for μ_e and σ_e in such an instance, it is possible to provide a basis for decision-making, using the likelihood ratio statistic along with a suitably

constructed hypothesis (reference 4). A result of this type is referred to as degenerate. It should be mentioned that, for a fixed population standard deviation, $\sigma \neq 0$, the probability of obtaining a degenerate result approaches 0 as the sample size becomes large. The proof of these statements is contained in reference 4 and is beyond the scope of this paper.

5. STATISTICAL PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATES. The preceding sections describe the method for selecting stress levels and the likelihood equations for calculating μ_e and σ_e. In order to make practical applications of the method, however, information is required regarding the statistical properties of the maximum likelihood estimates corresponding to various sample sizes. To facilitate obtaining this information, an extensive computer simulation was carried out using Monte Carlo techniques. In this way, hundreds of values of μ_e and σ_e could be obtained corresponding to hundreds of simulated experiments using random numbers for part strengths. Statistical summaries were then obtained and the variance of the estimates of the parameters empirically derived.

To perform the simulation, a standard interval, A = 1, B = 1, was chosen. The sampling of strengths was simulated by converting the sum of twelve two-digit numbers, constructed from a file of one millions random digits, to a random deviate with population mean μ and standard deviation, σ. Two populations were employed: $\mu = 0$ and $\sigma = 0..25$, and $\mu = 0.2$ and $\sigma = 0.1$. For each population, one hundred runs, each consisting of N samples, were made for N = 4 through 15, 20, 25, 30 and 35, with an additional four hundred runs for N = 15 and N = 30 for additional information on the distributions of the estimates. Finally, four runs of N = 900 were made for each population to empirically investigate the asymptotic convergence of (μ_e, σ_e) to (μ, σ).

For each set of 100 runs at a fixed sample size, the mean and variance were averaged separately and plotted as shown in Figures 3 and 4. Straight lines were then fitted to the data, recognizing that some spurious effects are introduced for small sample sizes (i. e., 15 or less) due to the discreteness of the admissible outcomes which are possible. (For sample size of N, only 2^N outcomes are possible corresponding to the 2^N possible configurations of 0 and 1.)

Based on the results of Figures 3 and 4, the variance of the estimates are given by:

VARIANCE OF THE MEAN

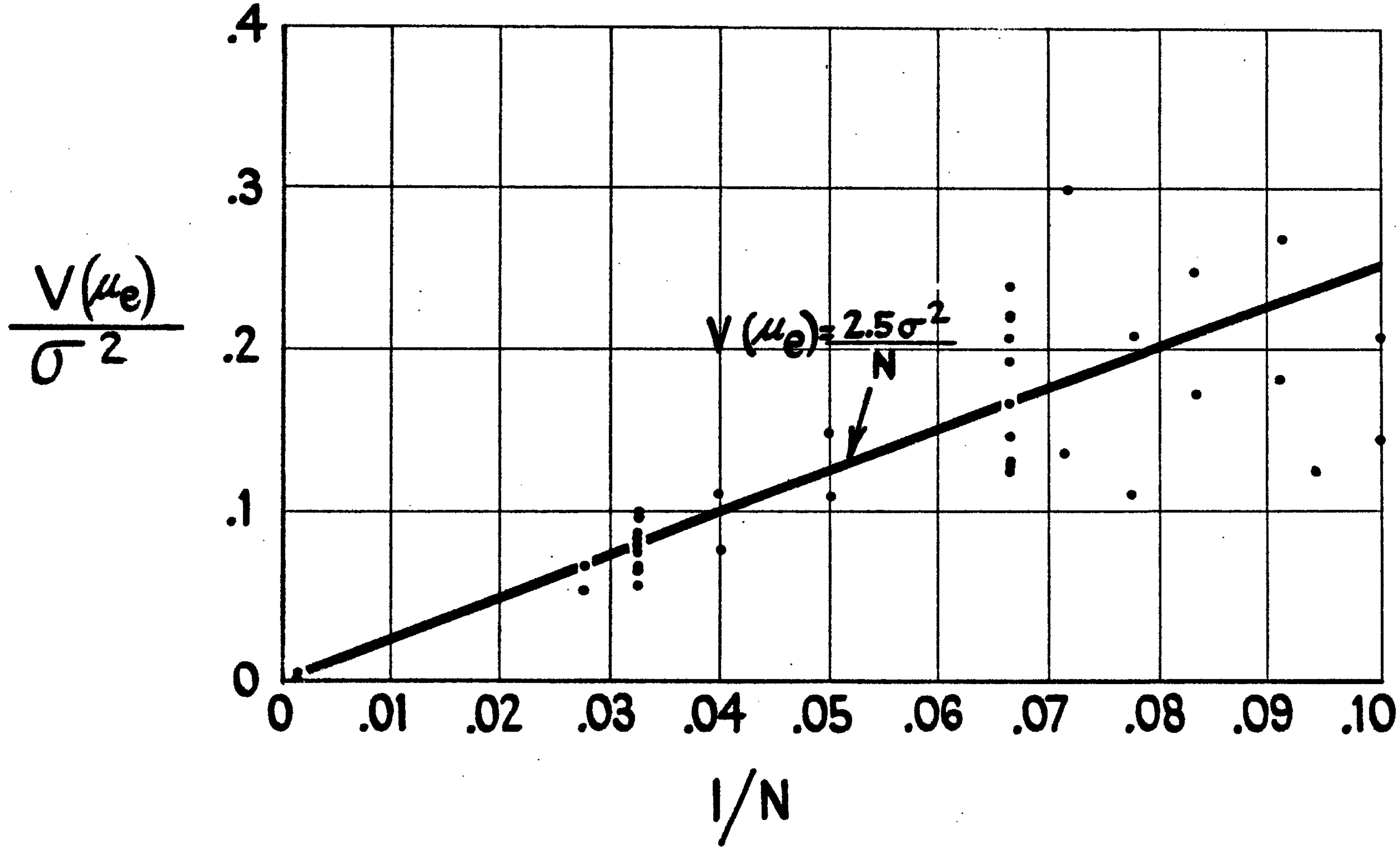


FIGURE 3.

VARIANCE OF THE STANDARD DEVIATION

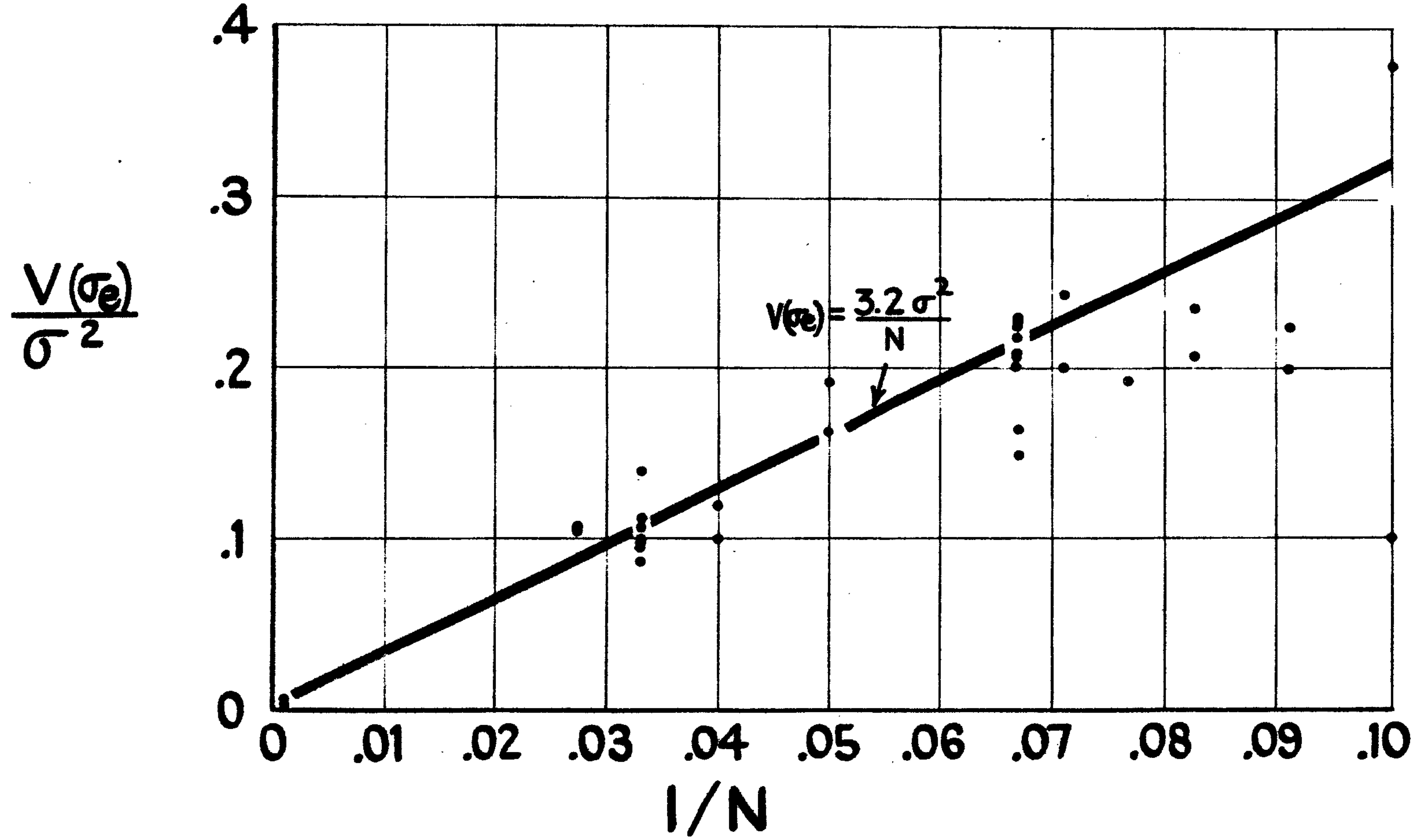


FIGURE 4.

$$V(\mu_e) = 2.5\sigma^2/N$$

$$V(\sigma_e) = 3.2\sigma^2/N$$

In order to use these formulas in practical applications, the unbiased estimate of σ, denoted $\hat{\sigma}$, is substituted for σ in the above formulas. The relationship between the unbiased estimate and the maximum likelihood estimate is given empirically in Figure 5. The variance of the unbiased estimate can be calculated using the relation

$$V(\hat{\sigma}) = \frac{V(\sigma_e)}{\beta^2} = \frac{3.2\hat{\sigma}^2}{N\beta^2} = \frac{3.2\sigma_e^2}{N\beta^4}$$

The above formula is sufficiently accurate for sample sizes on the order of 50 or greater, wherein the distribution of the estimate of the standard deviation approaches the normal distribution. For smaller size samples it was observed from the empirical study that $n\hat{\sigma}/\sigma$ approximately follows the chi-square distribution with n degress of freedom where n is given by

$$n = 0.625\beta^2 N \text{ (reference 4)}$$

where N is the sample size and β is given by Figure 5.

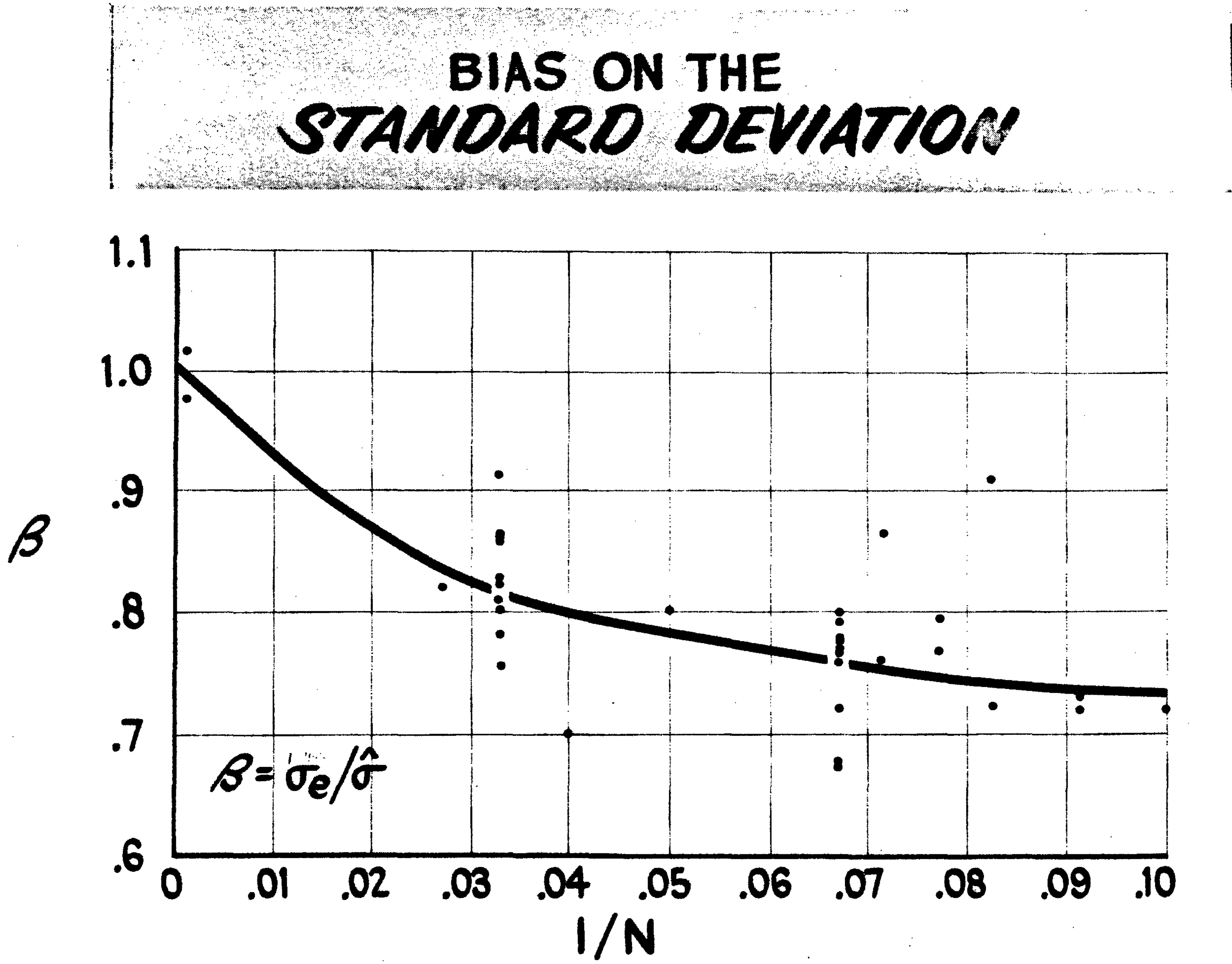

BIAS ON THE
STANDARD DEVIATION
1.1
1.0
.9
.8
.7
.6
β
0
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
1/N
$\beta = \sigma_e / \hat{\sigma}$

FIGURE 5.

REFERENCES

1. Lusser, R., Reliability Through Safety Margin, Research and Development Division, Army Rocket and Guided Missile Agency, Redstone Arsenal, Alabama, October 1958.

2. Dixon and Massey, Introduction to Statistical Analysis, McGraw-Hill 2nd Edition, 1957, Chapter 6.

3. Finney, D. J., Probit Analysis, Cambridge University Press

4. Langlie, H. J., A Reliability Test Method for "One-Shot" Items, Aeronutronic Division of Ford Motor Company Publication No. U-1792, August 10, 1962.

5. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, 6th printing, 1954, p.478.